从田野调查看语言的记录与保护

汪摇锋

(北京大学 中国语言学研究中心, 北京摇100871)

摘摇要: 在语言濒危情况日益严峻的情况下,对语言记录和保护显得更为迫切。本文从语言学田野调查的角度着重阐述了在语言记录和保护工作中应该注意的几个基本问题:在语言记录和保护中,语言材料的最基本要求是通过同音字表的检验;可以根据记录和保护的目的,选择合适的语言样本,根据需要呈现其中部分样本的;实际的语言记录和保护工作与理论假设之间是一种互动关系,相互制约又相互促进;由于双语在语言发展变化中的重要作用,应该对双语状况进行充分的描写。只有在这些基本问题上有清楚的认识,语言记录和保护才能更有成效。

关键词: 田野调查; 语言记录; 语言保护; 白语

中图分类号:H0**摇摇文献标识码:**A**摇摇文章编号:**1000-5919 (2013)03-0098-06

一

语言是文化最重要的物质形式之一,很早以来,语言的记录就受到重视。美国结构语言学派提出的发现程序就是为了记录语言,也为如何记录语言提供了基本的方法论基础。随着社会的发展,人们发现很多语言如物种一样开始消失,而且速度在加快,这就意味着语言多样性的丧失。“濒危语言”已经引起了广泛的关注,例如 2000 年 2 月在德国科隆召开了濒危语言学国际会议,将濒危语言分为 7 个等级,分别是:安全的语言;稳定但受到威胁的语言;受到侵蚀的语言;濒临危险的语言;严重危险的语言;濒临灭绝的语言;灭绝的语言。如今已经有几百种语言灭绝了,濒危的语言达千种。

应对语言濒危,通常有两种途径,即语言记录与语言保护。如 2010 年成立的世界语言记录与保护协会 (Consortium on Training in Language Documentation and Conservation)宣称其宗旨为:一方面进行语言记录,一方面促进各种语言的使用。语言记录着重对一个语言进行系统、完整的调查和记录;语言保护着重通过社会性手段来增加语言使用者。前者更具有语言学的本体价值,后者更具有语言使用的社会价值。

虽然着重点不同,但语言记录和语言保护关注的都是语言,因此有一些根本的问题必须共同面对。这些问题包括什么是一个语言系统中最基本的要素? 如何选取语言的样本才具有代表性? 记录和保护之前,需要什么样的理论假设? 双语与单语在语言发展中有什么样的位置? 这些基本的问题,也是语言田野调查必须面对的。

几乎所有的语言记录和保护都绕不开田野调查这一基础性工作。下文将从田野调查的角度来讨论语言的记录与保护的这几个基本问题。

二

几乎所有的语言记录和保护工作都会宣称对某个语言进行全面系统的保护。随着科学技术的

收稿日期:2013-03-18

作者简介:汪锋,男,湖北广水人,北京大学中文系副教授,中国语言学研究中心研究员。

基金项目:霍英东教育基金会“勒墨白族的语言及口传文化的数字化保护及基础理论研究”,项目批准号:131102;国家社会科学重大项目“中国有声语言及口传文化保护与传承的数字化方法研究和基础理论研究”,项目批准号:10&ZD125;教育部人文社会科学重点研究基地重大项目“基于系统语音对应的核心词分阶及建模研究”,项目批准号:11JJD740004。

进步,人们记录和保护语言的工具也发生了变化,从早期的纸和笔,发展到近现代的录音设备和录像设备,最近更是加上了各种生理采集设备,一套完整的数字化保护方法正在形成。

由于现代仪器可以记录多种多样的有关语言的信号,很多人开始觉得只要利用高分辨率的现代录音、录像等设备,就可以很方便地记录一个语言。由于几乎所有的物理和生理信号都高质量地采集到了,系统记录和保护的基础就具备了。这样的想法是一种误解。语言信号的物理属性和生理属性很重要,但却不是语言信号最本质的特点,语言信号最本质的特点是其功能属性,也就是作为交际系统,其承担的区别和表达意义的功能。因此,记录和保护语言的最低要求是完整记录一个语言的功能属性。

就语音而言,在语言学的田野调查中最重要的一项是提取音位。语言学教科书中都有对于提取音位的介绍。做过田野调查的语言学者都清楚,最开始进行记录时,使用国际音标记音,即,根据调查者的记音能力记录下被调查语言的词汇或者语句(不同的语言流派有不同的操作方式,有的学派主张一开始就记录自然语句,认为如果一开始就记录语词就带入了过多调查者的观念,可能调查到的不是“纯净”的对象)。记录到一定数量时,会遇到一些相似的音,也会产生提取音位的环境——最小对立(minimal pair),例如,汉语普通话中的“大[ta51]”和“榻[t^ha51]”,这时才可以提取两个音位/t/和/t^h/,如果一直没有遇到“榻[t^ha51]”这样的语素来跟“大[ta51]”构成在辅音上的对立,就不能肯定该语言中是否有两个对立的音位/t/和/t^h/。尽可能多地记录词汇,是田野调查者努力的一个方向。不同的语言词汇量的大小不一样,像汉语这样的语言,词汇量达到数万,而一些少数民族语言,词汇量可能不到三千。调查者发现很难再发现新词汇时,就要开始一项十分重要的工作,就是归纳音位。最小对立环境是提取音位的充分条件,但有些音从来都不对立,如,汉语普通话中的[æ]和[ɑ],这时就引入“互补相似”原则,即,如果分布环境是互补的,声音相似,则可以归纳为一个音位。这个原则的引入,也造成了音位归纳的灵活性。(例如,汉语普通话中的[æ]和[ɑ]这两个音是否相似,不同的学者有不同的看法。)因此,在田野调查中,最低的要求是对立的音位一定要区分出来。能保证这项最低要求的一个必要措施就是核对同音字表。

核对同音字表在传统语言调查中几乎是必备步骤,但近来受到科学技术的冲击,由于录音和录像的方便,短短几天就可以录完一个语言的词汇。一些调查者录完音后,就从田野离开,回去把录音转写为国际音标,以为记录工作就算完成了,或者自己仔细比对录音,来判定语音系统。很多调查者认为这是客观的操作,因为录音设备忠实地记录了语音。问题主要出在后来的转写上,音位系统恰恰是不能通过录音来忠实记录的。一个语言的音位系统是语言使用者对语音的感知范畴的划分,什么样的音是相同的,什么样的音是不同的,不同的语言有不同的分法,这也是语言多样性的一个体现。汉语普通话者对[æ]和[ɑ]的不同不在意,因为在他们的感知范畴中不区分二者,它们是一个音位/a/的条件变体,但英语者却很敏感,因为英语中它们是两个不同的音位,在感知范畴上严格区分。田野调查中一再强调的一个原则是“发音人永远是对的”。其根本原因就在于,对于语音异同的判断依赖于发音人的感知范畴。核对同音字表,就是让发音人来确定哪些音是相同的,哪些是不同的。这一基本工作必须依赖发音人。有田野调查经验的人都知道,经过这一核对过程,就会发现自己的有些记音过细,发音人不区分的,我们做了区分,有些记音过宽,发音人要区分的,我们没有区分出来。所谓语言记录和保护,自然要体现发音人在声音感知范畴上的性质,否则,记录和保护的只是调查者范畴感知下的语言,那就失去了语言记录和保护的本义。

一个语言的词汇在理论上说是有限的,也就是说,语音对比的环境是有限的。但这并不意味着,在田野调查中就一定能穷尽,因为有一些概念或者事物是被调查语言中独有的,非本族人难以知道,因此,在调查时就很难短时间内调查到,也就是说,系统完整的记录需要时间,并非一蹴而就的事情。

三

语言记录和保护计划一旦落实，马上就会遇到问题，记录和保护的对象如何确定？比如说，我们要记录和保护白语，如何确定工作对象？说白语的人口上百万，不可能都记录。这就是一个样本选择的问题。“样本”是相对于整体来说的，问题在于，这个整体是什么往往也并不清楚。仍然以白语为例，什么是白语？由于语言和民族是两个问题，不能简单地等同定义，不能说白族人说的就是白语。国家划分的白族支系，如勒墨和拉玛，其中有不少人认为自己的族群跟大理州的白族有很大的差距，在语言上也不容易沟通，对把自己族群归入白族之下有不同意见。在语言学上，语言与方言的区分也是见仁见智。汉语很多方言，比如，闽方言、粤方言等就被很多国外学者算作独立的语言。

从田野调查角度来看，确定这些工作对象是在调查之前要决定的。也就是说，调查对象是跟调查目的紧密相连的。调查的目的决定了对调查对象的选择。只要调查的对象能实现调查的目的，在样本选择上就是合适的，否则就不合适。以白语为例，很多研究只是取剑川金华镇的白语方言材料为据，就宣称白语的情况如何，或者断言白语和汉语在某个方面的异同，有的研究甚至不给出材料的来源。这就很容易误导读者，让人以为这是整个白语共同的普遍现象。如果以这些研究为基础，可能就会认为白语的记录已经不少了，在语言保护的评估中就会对白语的语言特点以及语言使用状况认知不足，造成以偏概全的状况。

值得注意的是，只要调查研究明确指出调查的目的和样本获取情况，就符合语言记录的基本要求，但调查的目的会发生变化，因此，在看到一个特定语言记录时要切合调查报告当时的目的，对样本获取等进行客观的评价。在白语的调查研究中，早期的《白语简志》（徐、赵 1984）以剑川金华镇白语方言为代表对白语进行了描述，最后还给出了剑川、大理、碧江三地的约 1000 多个词汇。这部简志对于人们了解白语起了很大的作用，功不可没。当时的调查目的是尽快获得对少数民族语言状况的一个基本了解，而不是对某个点的深入全面调查。根据作者的后记，可以看出调查集中于 1957 年、1959 年和 1960 年，其中后两次调查是补充性质的，“先后搜集了四十三个点的语言材料，归纳了近廿十个点的音位系统，进行了方言划分等研究工作”。这些基本调查工作服务于当时民族划分这一重大任务，最重要的是识别语言的异同，并选出标准点，创制民族语言文字。随着对语言状况了解的加深，对语言记录的要求也相应提高，现在再去调查白语方言，我们就不会满足于 1000 余词汇，而是希望挖掘更多具有白族特色的词汇；也不会满足于以松/紧的区别来命名白语发音上的特色，而是希望从其声学特性以及生理特性了解其根本性质。这些“不满足”就是从不同目的出发，对语言记录产生的新的样本要求。相信在将来，上面提到的这些希望会成为语言记录基本的要求，而很多新的希望会随着语言学的发展和科技的发展被提出来。从这个角度看，语言记录是一个动态发展的工作，语言保护也是，那么，自然地，样本选择也随之变化。

样本呈现也会随着语言记录和保护目的的变化而变化。样本呈现和样本选择并不是一回事。例如，《白语简志》的基础调查工作搜集了四十三个白语点的材料，但呈现在书稿中的却只有三个点，也就是，作者研究划分的三大白语方言的代表点。但是，今天看来，这三大白语方言的划分主要是根据共时差异。样本选择是根据记录或者保护目的进行的最初准备，样本呈现则是经过初步研究之后，把最能体现记录或者保护目的的样本发表出来。如果从了解白语方言的历史分化以及共时分布等细致的语言研究目的出发，四十三个方言点的材料就很令人期待。笔者从 2001 年开始对白语的发展变化以及历史来源感兴趣，如果能看到这四十三个点的材料，研究选点以及进行历史比较就更有基础，更有把握。由于笔者关注的是如何从白语方言比较来重构原始白语，并进一步与汉语和彝语等进行比较，在选择样本时就尽力寻找差别大的白语方言，根据前人的研究以及白语的地理分布，大约调查了 20 种方言，最后从中选择了 9 个方言作为重构原始白语的基础（Wang 2006；汪锋 2012），是由于这 9 个方言能涵盖所调查过的 20 种方言表现出的语音对应，也就是

说,不会漏掉一条语音对应。这样的样本呈现是根据笔者历史比较的目的。

语言学家的调查记录中,发音人往往是一个人,但在论文或者研究报告中,往往称这是某语言的状况,谨慎一些的会说是某地的语言状况如何。这个“某地”在不同的语言学家看来也有宽严之分,宽泛的可以大到县市,严格的小到自然村。因此,很多人对此现状很不满意,以为语言学家不够严谨,凭一个发音人的情况怎么能推断整个地区的语言状况呢?样本量也太小了吧?语言学家也很为难,要知道,调查一个发音人,进行详尽的记录,就要费不少时间和精力,如果按照社会调查取样的要求,覆盖各种社会因素,语言学家的精力、时间和经费都不够。从这个矛盾中,可以看出语言调查和记录的特殊性,与一般的社会调查不同,很难进行大样本的全面记录。同时,从语言学的角度来看,这样的大样本调查和记录或许也没有必要。语言是沟通的工具,一个言语社团,比如,一个自然村,几百人的语言之间相差不多,不同代之间可能在某些方面有变异,从而构成所谓“有序异质”的系统(Weinreich et al, 1968)。一方面,我们要注意到语言社团中可能存在变异,另一方面,我们也不能过于夸大变异的程度。因此,在调查记录的时候以一、两个发音人为主,同时注意其他人是否表现差异,在哪些方面有差异,可以针对差异进行专项调查,这时候就可以按照社会调查的取样办法来操作。至于说,一个发音人的语言能代表多大范围的言语社团,就不能一概而论,有些地方是“乡隔一丛草,讲话不知道”。有些地方是沿途几百里,一种语言畅通无阻。因此,语言学家记录的时候要准确地给出发音人详细信息即达到基础要求,至于其能代表多大范围的言语,是需要进一步调查研究的事情,同样不能一蹴而就。

四

语言记录和保护常常被看做是客观的记录和保护,仿佛语言是一个客观存在,记录者用客观的手段,比如录音或者录像,就可以把语言记录下来。事实上,没有“一个”语言摆在田野等着我们去记录,语言呈现在言语社团成员的言语中,而这些言语是有序异质的,如何从这些言语切入进去,进行语言记录和语言保护,不同学者有不同的看法。而且,对于任何一个学科来说,所谓零度客观的观察几乎没有,只要选择如何切入,就意味着选择了某种理论假设。

以汉语方言的调查记录为例,是否选择《方言调查字表》作为辅助,就可以说明调查者的倾向。使用《方言调查字表》的调查者主要是想获得方言的“字音”,以与其他方言和历史上的音韵分合进行比较,高本汉可谓其开创者,后来,赵元任等继续这一模式,并以《方言调查字表》的形式将之精炼化和可操作化,但后来一些调查者认为方言自然的口语也非常重要,“字音”掩盖了很多语言发展变化的信息,也有一些缺陷,提倡调查方言的口语词,以罗杰瑞等为代表。① 高晓虹详细分析了这两种调查切入方式产生的原因以及各自的限制,可参看。② 从中,我们可以清楚看出,调查什么,不调查什么;重点调查什么;如何记录调查结果等细节无不体现出调查者在理论上所做的假定,不带理论假设或者无特定目的的纯粹语言记录和保护是不存在的。

不同语言观和不同的语言兴趣会造成不同的语言记录方式,不同的语言记录方式会切合不同的记录目的。例如,Labov 对语言的变异感兴趣,他要观察言语社团中 -r 的发音是否存在差异,就设计了一个专门的调查,到百货大楼去调查顾客(Labov 1966)。没有语言变异观的语言调查者不会注意到这一现象。中国少数民族语言的调查早期很受汉语的影响,从“中国少数民族语言简志丛书”就可以看出来,大部分是从少数民族语言中找汉语语言系统的对应物。现在的少数民族语言调查和记录就丰富多彩了,有类型学框架下的各种参考语法,有专注语音研究的声学-生理研究等。以我研究白语的情况为例,刚开始研究时,关注的主要是白语方言之间的音韵对应及其历史发展过程,因此,在调查记录的时候主要关注

① 罗杰瑞:《汉语方言田野调查与音韵学》,《北京大学学报》2007 年第 2 期。

② 高晓虹:《汉语调查方法的思考》,《语言学论丛》第 43 辑,北京:商务印书馆 2011 年版,第 164—176 页。

词汇和音系。在这方面,值得一提的是《白语简志》,作者之一徐琳本人是白语母语者 (尽管不是剑川金华镇的,而是洱源县乔后镇),在音系的感知上有天然的优势,该简志对于音位系统的提取也基本根据严格的最小对立。由于早期对于语言发声类型认识还不够清楚,对特殊发声类型和普通发声类型的对立以松紧范畴来标示,并与音高结合,整合进声调系统,符合母语者的范畴认知,在语言学上是精彩的处理案例。但随着调查研究的深入,我对白语中所谓“紧音”产生兴趣,开始研究其性质,这时候声学分析以及嗓音的生理信号收集等方式就随之带入,否则,调查和研究几乎难以进行,通过研究,意识到白语的“紧音”属于特殊发声类型,各个方言有不同的嗓音机制,其变异性和混合性在中国少数民族语言中是难得一见的,这也就更突出了对其进行记录和保护的价值。①

五

在语言记录时,常见到一种说法,也就是,尽量调查单语者,因为单语者更纯粹。其中隐含的意思是双语者的语言会受到混合或者干扰。这里就有一个未经证明的假定,即,语言越少“混合”越有保护价值。如果继续追问,为什么越“纯粹”的语言越值得保护,可能的回答是,越纯粹,则受他语影响越少,则保留早期的面貌越完整。可见,这里的保护价值主要是为了追溯历史的方便。

保护语言,并不单纯为了追溯历史。语言保护的价值可以体现在更多方面,其中一个最重要的方面是多样性。而语言多样性的造成跟语言接触有密切的关系,一方面在语言接触中,一些弱势语言会被强势语言替换,造成语言多样性的减少,另一方面,很多语言现象是在语言接触中造成的,甚至,很多新的语言也是在语言接触中产生的,不同的地域,语言接触程度和方式不同,也会造成语言发展模式的各种精彩变异,这是人类文化中十分值得珍视的部分。假定语言之间都不接触,都自我发展,一直保持单语状态,语言自身也会发生变化,但变化速度缓慢,生成新的语言现象的概率也不大。注意到这后一方面,就要仔细检视语言接触与语言濒危的关系,而不是一味反对语言接触。②

在语言记录和语言保护的过程中,我们发现双语其实比单语更有活力,以白语为例,白族很早就开始了白—汉双语状态,直到今天,差不多百分之八九十的白语者,同时也是汉语者,因此,白语延续发展至今 (至少从唐代就开始了)至少已经有千年以上。现在的白语不仅方言众多,语言现象丰富,而且使用人口众多,丝毫不“濒危”。与此相比,目前处于濒危状态的少数民族语言,很多都是单语状态时间保持太长,对强势的汉语影响没有缓冲,就直接被替换了。这提醒我们在语言记录和保护时,要充分注意到双语的缓冲作用。

在田野调查中,双语者作为发音人可以为语言记录提供很多便利,不仅在于时间节省方面,更重要的是,双语者能在两种语言中转换,对概念或者语音进行系统的匹配与对应。许多概念,如果从调查者的角度出发,很难向单语者问清楚,一方面调查者难以设定一个好的环境;另一方面,就算调查者习得了一些被调查语言的语句,但对整个概念语义系统的理解还需要很长时间的摸索,如果不生活在其中,是难以表达清楚的,这也是为什么很多学者主张“浸入式调查”的原因之一。③

六

语言记录和语言保护是人类珍视语言这一数万年来文化结晶的重要行动,但在付诸实施之前,应该对其中的基本问题有清楚的认识。本文再次从语言学田野调查的角度出发重申了对一些基本问题的看法,包括:1. 无论采录设备如何进步,以同音字表形式验证发音人的感知范畴都是必要

① 汪锋:《白语方言中特殊发声类型的来源与发展》,《汉藏语学报》第 1 期,北京:商务印书馆 2007 年版,第 162—170 页。

② 陈保亚:《从接触看濒危方言、濒危特征和濒危机制》,《长江学术》2006 年第 1 期。

③ 罗仁地:《语言类型学/功能语言学派视野下的语言学田野调查》,《语言学论丛》第 36 辑,北京:商务印书馆 2007 年版,第 42—56 页。

的;2. 语言样本的选择和呈现都与记录和保护的目的有关,没有一成不变的套路;3. 没有纯粹客观的语言记录和保护,理论假设一方面制约着记录和保护的方式,另一方面在记录和保护中得到发展;4. 语言并不是越纯净就越有价值,双语在语言发展中起着非常重要的作用,在记录和保护工作中应该加以重视。我们认为,如果对这些基本问题没有清楚的认识,语言记录和保护的效果会大打折扣。

(感谢杨海潮、张静芬、刘文、陈卓娅指正多处。)

参考文献:

高本汉 (1940/2003),《中国音韵学研究》,赵元任、罗常培、李方桂合译,北京:商务印书馆。

汪锋 (2012),《语言接触与语言比较——以白语为例》,北京:商务印书馆。

徐琳、赵衍荪 (1984),《白语简志》,北京:民族出版社。

Alexandra Y. Aikhenvald (2008),《语言学田野调查·代序》,吴君如译,戴庆厦、罗仁地、汪锋主编:《到田野去——语言学田野调查的方法与实践》,北京:民族出版社,第 1—14 页。

Labov, W. (1966). *The Social Stratification of /r/ in New York City Department Stores*. In N. Coupland, & A. Jaworski, A. (Eds.) Sociolinguistics: A Reader. (pp. 168 – 178), New York: St Martin's Press.

Wang, Feng (2006). *Comparison of Languages in Contact-The Case of Bai and the Distillation Method*, Taibei: Academia Sinica.

Wang, Feng (2011). *Phonation types in the Jianchuan Bai.* Chinese University of Hong Kong. Jan 12. 2011.

Weinreich, U., Labov, W and Herzog, M. I, 1968, *Empirical Foundations for a Theory of Language Change*, University of Texas Press.

Language Documentation and Conservation From the Perspective of Linguistic Fieldwork

Wang Feng

(*Research Center for Chinese Linguistics*, *Peking University*, *Beijing*摇100871, *China*)

Abstract: Language documentation and conservation are an important action to cherish languages as cultural fruits of tens of millennia. This paper restated opinions on several fundamental issues in language documentation and conservation from the perspective of linguistic fieldwork. Firstly, it is always necessary to examine and confirm the categories of the informants' perception by means of homophone lists, no matter how advanced the recording equipment might be. Secondly, the selection and representation of language samples are closely related to the purpose of language documentation and conservation, without a never-changing method. Thirdly, there is no absolutely objective language documentation and conservation. Theoretical hypotheses may limit the means of documentation and conservation on the one hand, and they may also be improved during the process of documentation and conservation on the other hand. Fourthly, it is not the case that the less interfered with the language is, the more valuable it is. Bilingualism plays a very important role in the language development, which should be paid much attention to in the documentation and conservation.

Key words: linguistic fieldwork, language documentation, language conservation, *Bai*

(责任编辑摇摇何摇方)